



PCT09

## RAW SEQUENCE LISTING

DATE: 06/05/2002

PATENT APPLICATION: US/09/980,862

TIME: 17:29:50

Input Set : A:\29029101.app

Output Set: N:\CRF3\06052002\I980862.raw

p.6

ENTERED

3 <110> APPLICANT: SCHLEHUBER, STEFFEN  
 5 <120> TITLE OF INVENTION: MUTEINS OF THE BILIN-BINDING PROTEIN  
 7 <130> FILE REFERENCE: 029029/0101  
 9 <140> CURRENT APPLICATION NUMBER: 09/980,862  
 10 <141> CURRENT FILING DATE: 2002-05-06  
 12 <150> PRIOR APPLICATION NUMBER: DE 199 26 068.0  
 13 <151> PRIOR FILING DATE: 1999-06-08  
 15 <160> NUMBER OF SEQ ID NOS: 27  
 17 <170> SOFTWARE: PatentIn Ver. 2.1  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 1219  
 21 <212> TYPE: DNA  
 22 <213> ORGANISM: Artificial Sequence  
 24 <220> FEATURE:  
 25 <223> OTHER INFORMATION: Description of Artificial Sequence: pBBP20  
 26 nucleic acid sequence  
 28 <220> FEATURE:  
 29 <221> NAME/KEY: sig\_peptide  
 30 <222> LOCATION: (22)..(84)  
 32 <220> FEATURE:  
 33 <221> NAME/KEY: mat\_peptide  
 34 <222> LOCATION: (85)..(1209)  
 35 <223> OTHER INFORMATION: fusion protein of bilin-binding protein, Strep-tag II  
 36 and fragment of phage coat protein pIII  
 38 <220> FEATURE:  
 39 <221> NAME/KEY: CDS  
 40 <222> LOCATION: (85)..(606)  
 41 <223> OTHER INFORMATION: mature bilin-binding protein  
 43 <220> FEATURE:  
 44 <221> NAME/KEY: CDS  
 45 <222> LOCATION: (607)..(636)  
 46 <223> OTHER INFORMATION: Strep-tag II-affinity tag  
 48 <220> FEATURE:  
 49 <221> NAME/KEY: misc\_feature  
 50 <222> LOCATION: (637)..(639)  
 51 <223> OTHER INFORMATION: amber stop codon  
 53 <220> FEATURE:  
 54 <221> NAME/KEY: CDS  
 55 <222> LOCATION: (640)..(1209)  
 56 <223> OTHER INFORMATION: amino acids 217-406 of coat protein pIII  
 58 <400> SEQUENCE: 1  
 59 tctagtttaac gagggcaaaa a atg aaa aag aca gct atc gcg att gca gtg 51  
 60 Met Lys Lys Thr Ala Ile Ala Ile Ala Val

## RAW SEQUENCE LISTING

DATE: 06/05/2002

PATENT APPLICATION: US/09/980,862

TIME: 17:29:50

Input Set : A:\29029101.app

Output Set: N:\CRF3\06052002\I980862.raw

```

61                                     -20                                     -15
63 gca ctg gct ggt ttc gct acc gta gcg cag gcc gac gtg tac cac gac 99
64 Ala Leu Ala Gly Phe Ala Thr Val Ala Gln Ala Asp Val Tyr His Asp
65      -10                                     -5                                     -1 1 5
67 ggt gcc tgt ccc gaa gtc aag cca gtc gac aac ttc gac tgg tcc cag 147
68 Gly Ala Cys Pro Glu Val Lys Pro Val Asp Asn Phe Asp Trp Ser Gln
69      10                                     15                                     20
71 tac cat ggt aaa tgg tgg gaa gtc gcc aaa tac ccc aac tca gtt gag 195
72 Tyr His Gly Lys Trp Trp Glu Val Ala Lys Tyr Pro Asn Ser Val Glu
73      25                                     30                                     35
75 aag tac gga aag tgc gga tgg gct gag tac act cct gaa ggc aag agt 243
76 Lys Tyr Gly Lys Cys Gly Trp Ala Glu Tyr Thr Pro Glu Gly Lys Ser
77      40                                     45                                     50
79 gtc aaa gtt tgc aac tac cac gta atc cac ggc aag gaa tac ttt att 291
80 Val Lys Val Ser Asn Tyr His Val Ile His Gly Lys Glu Tyr Phe Ile
81      55                                     60                                     65
83 gaa gga act gcc tac cca gtt ggt gac tcc aag att gga aag atc tac 339
84 Glu Gly Thr Ala Tyr Pro Val Gly Asp Ser Lys Ile Gly Lys Ile Tyr
85      70                                     75                                     80                                     85
87 cac agc ctg act tac gga ggt gtc acc aag gag aac gta ttc aac gta 387
88 His Ser Leu Thr Tyr Gly Gly Val Thr Lys Glu Asn Val Phe Asn Val
89      90                                     95                                     100
91 ctc tcc act gac aac aag aac tac atc atc gga tac tac tgc aaa tac 435
92 Leu Ser Thr Asp Asn Lys Asn Tyr Ile Ile Gly Tyr Tyr Cys Lys Tyr
93      105                                     110                                     115
95 gac gag gac aag aag gga cac caa gac ttc gtc tgg gtg ctc tcc aga 483
96 Asp Glu Asp Lys Lys Gly His Gln Asp Phe Val Trp Val Leu Ser Arg
97      120                                     125                                     130
99 agc atg gtc ctt act ggt gaa gcc aag acc gct gtc gag aac tac ctt 531
100 Ser Met Val Leu Thr Gly Glu Ala Lys Thr Ala Val Glu Asn Tyr Leu
101      135                                     140                                     145
103 atc ggc tcc cca gta gtc gac tcc cag aaa ctg gta tac agt gac ttc 579
104 Ile Gly Ser Pro Val Val Asp Ser Gln Lys Leu Val Tyr Ser Asp Phe
105 150                                     155                                     160                                     165
107 tct gaa gcc gcc tgc aag gtc aac aat agc aac tgg tct cac ccg cag 627
108 Ser Glu Ala Ala Cys Lys Val Asn Asn Ser Asn Trp Ser His Pro Gln
109      170                                     175                                     180
111 ttc gaa aaa tag gct ggc ggc ggc tct ggt ggt ggt tct ggc ggc ggc 675
112 Phe Glu Lys Gln Ala Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly
113      185                                     190                                     195
115 tct gag ggt ggt ggc tct gag ggt ggc ggt tct gag ggt ggc ggc tct 723
116 Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser
117      200                                     205                                     210
119 gag gga ggc ggt tcc ggt ggt ggc tct ggt tcc ggt gat ttt gat tat 771
120 Glu Gly Gly Gly Ser Gly Gly Ser Gly Ser Gly Asp Phe Asp Tyr
121      215                                     220                                     225
123 gaa aag atg gca aac gct aat aag ggg gct atg acc gaa aat gcc gat 819
124 Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp
125 230                                     235                                     240                                     245

```

## RAW SEQUENCE LISTING

DATE: 06/05/2002

PATENT APPLICATION: US/09/980,862

TIME: 17:29:50

Input Set : A:\29029101.app

Output Set: N:\CRF3\06052002\I980862.raw

```

127 gaa aac gcg cta cag tct gac gct aaa ggc aaa ctt gat tct gtc gct 867
128 Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala
129                250                255                260
131 act gat tac ggt gct gct atc gat ggt ttc att ggt gac gtt tcc ggc 915
132 Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly
133                265                270                275
135 ctt gct aat ggt aat ggt gct act ggt gat ttt gct ggc tct aat tcc 963
136 Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser
137                280                285                290
139 caa atg gct caa gtc ggt gac ggt gat aat tca cct tta atg aat aat 1011
140 Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn Asn
141                295                300                305
143 ttc cgt caa tat tta cct tcc ctc cct caa tgc gtt gaa tgt cgc cct 1059
144 Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg Pro
145 310                315                320                325
147 ttt gtc ttt ggc gct ggt aaa cca tat gaa ttt tct att gat tgt gac 1107
148 Phe Val Phe Gly Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp
149                330                335                340
151 aaa ata aac tta ttc cgt ggt gtc ttt gcg ttt ctt tta tat gtt gcc 1155
152 Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala
153                345                350                355
155 acc ttt atg tat gta ttt tct acg ttt gct aac ata ctg cgt aat aag 1203
156 Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys
157                360                365                370
159 gag tct taataagctt 1219
160 Glu Ser
161 375
164 <210> SEQ ID NO: 2
165 <211> LENGTH: 64
166 <212> TYPE: DNA
167 <213> ORGANISM: Artificial Sequence
169 <220> FEATURE:
170 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
172 <220> FEATURE:
173 <221> NAME/KEY: modified_base
174 <222> LOCATION: (35)..(36)
175 <223> OTHER INFORMATION: a, t, c, g, other or unknown
177 <220> FEATURE:
178 <221> NAME/KEY: modified_base
179 <222> LOCATION: (38)
180 <223> OTHER INFORMATION: a, t, c, g, other or unknown
182 <220> FEATURE:
183 <221> NAME/KEY: modified_base
184 <222> LOCATION: (41)..(42)
185 <223> OTHER INFORMATION: a, t, c, g, other or unknown
187 <220> FEATURE:
188 <221> NAME/KEY: modified_base
189 <222> LOCATION: (44)..(45)
190 <223> OTHER INFORMATION: a, t, c, g, other or unknown

```

## RAW SEQUENCE LISTING

DATE: 06/05/2002

PATENT APPLICATION: US/09/980,862

TIME: 17:29:50

Input Set : A:\29029101.app

Output Set: N:\CRF3\06052002\I980862.raw

192 <400> SEQUENCE: 2  
 193 ccatggtaaa tgggtgggaag tcgccaaata ccccnknms nnsnnkaagt acggaaagtg 60  
 194 cgga 64  
 197 <210> SEQ ID NO: 3  
 198 <211> LENGTH: 71  
 199 <212> TYPE: DNA  
 200 <213> ORGANISM: Artificial Sequence  
 202 <220> FEATURE:  
 203 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer  
 205 <220> FEATURE:  
 206 <221> NAME/KEY: modified\_base  
 207 <222> LOCATION: (19)..(20) /  
 208 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 210 <220> FEATURE:  
 211 <221> NAME/KEY: modified\_base  
 212 <222> LOCATION: (46)..(47) /  
 213 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 215 <220> FEATURE:  
 216 <221> NAME/KEY: modified\_base  
 217 <222> LOCATION: (52)..(53) /  
 218 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 220 <400> SEQUENCE: 3  
 221 gggtaggcgg tacctcsnn aaagtattcc ttgccgtgga ttacmngta snncgaaact 60  
 222 ttgacactct t 71  
 225 <210> SEQ ID NO: 4  
 226 <211> LENGTH: 74  
 227 <212> TYPE: DNA  
 228 <213> ORGANISM: Artificial Sequence  
 230 <220> FEATURE:  
 231 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer  
 233 <220> FEATURE:  
 234 <221> NAME/KEY: modified\_base  
 235 <222> LOCATION: (27)..(28) /  
 236 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 238 <220> FEATURE:  
 239 <221> NAME/KEY: modified\_base  
 240 <222> LOCATION: (33)..(34) /  
 241 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 243 <220> FEATURE:  
 244 <221> NAME/KEY: modified\_base  
 245 <222> LOCATION: (42)..(43) /  
 246 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 248 <220> FEATURE:  
 249 <221> NAME/KEY: modified\_base /  
 250 <222> LOCATION: (54)..(55)  
 251 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 253 <400> SEQUENCE: 4  
 254 ccaagattgg aaagatctac cacagcnnsa ctnnkggagg tnsaccvvs gagnnkgtat 60  
 255 tcaacgtact ctcc 74

## RAW SEQUENCE LISTING

DATE: 06/05/2002

PATENT APPLICATION: US/09/980,862

TIME: 17:29:50

Input Set : A:\29029101.app

Output Set: N:\CRF3\06052002\I980862.raw

```

258 <210> SEQ ID NO: 5
259 <211> LENGTH: 78
260 <212> TYPE: DNA
261 <213> ORGANISM: Artificial Sequence
263 <220> FEATURE:
264 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
266 <220> FEATURE:
267 <221> NAME/KEY: modified_base
268 <222> LOCATION: (20)..(21)
269 <223> OTHER INFORMATION: a, t, c, g, other or unknown
271 <220> FEATURE:
272 <221> NAME/KEY: modified_base
273 <222> LOCATION: (26)..(27)
274 <223> OTHER INFORMATION: a, t, c, g, other or unknown
276 <220> FEATURE:
277 <221> NAME/KEY: modified_base
278 <222> LOCATION: (53)..(54)
279 <223> OTHER INFORMATION: a, t, c, g, other or unknown
281 <220> FEATURE:
282 <221> NAME/KEY: modified_base
283 <222> LOCATION: (59)..(60)
284 <223> OTHER INFORMATION: a, t, c, g, other or unknown
286 <400> SEQUENCE: 5
> 287 tctggagagc acccagacmn ngtcsmngtg tcccttcttg tcctcgtcgt asnngcamnn 60
288 gtatccgatg atgtagtt 78
291 <210> SEQ ID NO: 6
292 <211> LENGTH: 36
293 <212> TYPE: DNA
294 <213> ORGANISM: Artificial Sequence
296 <220> FEATURE:
297 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
299 <400> SEQUENCE: 6
300 cttcgactgg tcccagtacc atggtaaagtg gtggga 36
303 <210> SEQ ID NO: 7
304 <211> LENGTH: 37
305 <212> TYPE: DNA
306 <213> ORGANISM: Artificial Sequence
308 <220> FEATURE:
309 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
311 <400> SEQUENCE: 7
312 caccagtaag gaccatgctt ctggagagca cccagac 37
315 <210> SEQ ID NO: 8
316 <211> LENGTH: 46
317 <212> TYPE: DNA
318 <213> ORGANISM: Artificial Sequence
320 <220> FEATURE:
321 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
322 oligodeoxynucleotide
324 <400> SEQUENCE: 8

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/980,862

DATE: 06/05/2002  
TIME: 17:29:51

Input Set : A:\29029101.app  
Output Set: N:\CRF3\06052002\I980862.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; N Pos. 35, 36, 38, 41, 42, 44, 45  
Seq#:3; N Pos. 19, 20, 46, 47, 52, 53  
Seq#:4; N Pos. 27, 28, 33, 34, 42, 43, 54, 55  
Seq#:5; N Pos. 20, 21, 26, 27, 53, 54, 59, 60  
Seq#:13; N Pos. 29, 30, 38, 39, 47, 48, 50, 51, 53, 54, 56, 57

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:16; Line(s) 881